View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Research Papers in Economics

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

Working Paper 8906

DO WAGE DIFFERENCES AMONG EMPLOYERS LAST?

by Erica L. Groshen

Erica L. Groshen is an economist at the Federal Reserve Bank of Cleveland. The author thanks Richard Freeman, John Dunlop, Katherine Bradbury, John Bound, and seminar participants at Boston and Stanford Universities for their valuable comments. Partial funding from Social Sciences Research Council Grant No. SS-25-83-31 is gratefully acknowledged.

Working papers of the Federal Reserve Bank of Cleveland are preliminary materials circulated to stimulate discussion and critical comment. The views stated herein are those of the author and not necessarily those of the Federal Reserve Bank of Cleveland or of the Board of Governors of the Federal Reserve System.

June 1989

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

ABSTRACT

Recent interest in efficiency wage and insider/outsider models of wage determination has drawn attention to employer-based wage differences. *Alternatively*, these differences may simply reflect temporary, random errors by wage-setters. This paper provides strong evidence against the possibility that employer wage variations are temporary or randm, along with additional verification of the existence of substantial employer wage differences within and between industries.

The variance of wages is analyzed in a unique data set: wages paid to individual workers in selected blue- and white-collar occupations from a six-year panel of employers within a single standard metropolitan statistical area. The most conservative estimate of establishment wage differentials in this sample (controlling for very detailed job classification) yields a standard deviation of approximately 12 percent within industry, or 18 percent, including interindustry differentials. Wage differences among employers are *shown* to be virtually stationary over time and related to establishment size, but not consistently to changes in establishment employment.

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

I. Introduction

The existence of large employer-based wage differences among apparently equivalent workers is often taken as supporting evidence for the existence of efficiency wages or implicit profit-sharing (see Dickens and Katz [1987], for example).[1] The two main alternative hypotheses that have been explored are sorting by worker-quality and by compensating differentials, neither of which has found strong support in statistical tests. This paper tests a third alternative, whether wage differences among employers are the result of random, temporary errors.

If employer differentials are the result of errors, the efficiency of the labor market may be enhanced by their elimination, perhaps through government subsidies of information gathering and dissemination. On the other hand, if these differentials are efficient wages or profit-shares, they may be appropriate second-best solutions to monitoring or agency problems endemic to the labor market, but have implications for other policy, such as trade or antidiscrimination policy, as demonstrated in Bulow and Summers (1986), or for macroeconomic policy, as shown in Weitzman (1986).

Efficiency wage arguments posit causality between workers' wages and on-the-job productivity (Yellen [1984], Stiglitz [1984]). Thus, some employers may maximize profits by paying a differential above the market-clearing wage, if resulting increments in productivity exceed costs of the differential. At least five sources of increased productivity have been modeled: reduced monitoring (or shirking) costs (for example, Bulow and Summers [1986]), decreased turnover (Salop [1979]), sociological considerations (Akerlof [1982]), market insulation, and corporate consistency (Doeringer and Piore [1971]).

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

In contrast, implicit profit-sharing models of wage variation (also called insider/outsider, rent-sharing, and bargaining models) assume the existence of variations in firms' rents and in employees' bargaining power (or agency costs). These conditions introduce the possibility of rent-capture by employees, although the models differ in the identity of agents and enforcement mechanisms. The players are clearest in the case of unionism; otherwise, the workers' bargaining agent is not obvious, although economists have long noted the existence of informal organization by nonunion workers (Dunlop [1957]), including union-threat effect versions (Dickens [1986]) and managerial capitalism/agency cost versions (Aoki [1984]).

This paper focuses on the alternative explanation that wage differences among employers simply reflect random errors by wage-setters. Seminal articles by Stigler (1962) and Rothschild and Stiglitz (1976) launched a family of pure information models that use costly job search to explain wage dispersion. Expensive job *search* allows the market to sustain a range of wages because a worker's gain from further search becomes uncertain, rather than a known quantity. While mean wages for a particular type of worker are *equal* to the worker's marginal product, the costs of information introduce an error term with a variance that is a positive function of the search and mobility costs for workers or employers. Thus, if employers adjust all workers' wages in tandem, errors may be correlated across occupations for an employer.

Most previous empirical studies of interemployer wage differentials have focused on national interindustry differences.[2] Because of data limitations, these studies have been unable to control well for local labor market conditions or detailed occupation, to compare differentials between industries to those within industry, or to investigate the stability of employer differentials over time.

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

This paper provides new insight into establishment-based wage variation, using a unique data set prepared for the author by the U.S. Bureau of Labor Statistics. The wages of nonsupervisory white- and blue-collar workers in one city are examined to see whether employer differentials exist within a single labor market, whether they are stable over the course of six years, and whether they are associated with *growth* or shrinkage of the establishment. Wage variation between industries is also compared to that within industry. In addition, the results are compared to those in the Current Population Survey in order to estimate the importance of interemployer wage variation as a source of wage variation in the economy as a whole.

The results cast light on the nature of wage differences among employers and on the plausibility of other proposed sources of wage variation by employer. A number of previous studies find it unlikely that employer differentials arise from systematic sorting of workers by measured or unmeasured ability within occupation.[3] Even stronger empirical evidence tends to refute the hypothesis that wage differences among employers compensate for establishmentwide variations in working conditions.[4] This paper provides evidence of substantial wage differences among employers within a single city. This finding greatly reduces the possibility that regionwide compensating differentials for cost of living are the main source of employer differentials.[5]

The major contribution of this paper is the finding that interemployer wage differences, and rankings of employers by wage, are virtually stationary over six years. This result eliminates random variations (generated or perpetuated by costly information) as a likely s o u .of employer differentials. The persistence of establishment wage differentials is consistent with earlier

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

findings that employer wage differences are associated with measurable characteristics of employers, such as establishment size and product market (Groshen [1988b]).

Process of elimination leaves the door open for the two provocative types of models of employer wage variation (efficiency wages and rent-sharing) that have generated considerable interest. The conclusion identifies several key characteristics of interemployer wage differentials that need to be accounted for in any version of these models invoked.

II. The Data

The data used in this study are a unique set compiled for the author by the U.S. Bureau of Labor Statistics, from Area Occupational Wage Surveys (AWS) for a single metropolitan statistical area (MSA) over the course of six years. The variables include the wage, sex, occupation, and establishment identifier of individual workers in nonsupervisory positions. Wages are the straight-time hourly wages (no overtime or shift premia included) of hourly workers, and the average hourly earnings of incentive workers. Although confidentiality restrictions prohibit the release of employers' names, the data include unique establishment identifiers and two plant characteristics: size class and two-digit Standard Industry Classification (SIC) code.

This survey has the following advantages: it allows control for MSA, it includes many different industries, and it is longitudinal in establishments. In addition, the surveys cover a broad mix of occupations: white- and blue-collar, professional, skilled, and unskilled. The occupations surveyed belong to four major groups: clerical/office workers, professional personnel, custodial/material-movement workers, and maintenance/toolroom/powerplant occu-

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

pations. (Appendix A presents a complete list of the occupations covered in the survey.)

An important feature of these data is specificity of the occupation definitions, which are actually job classifications and are more detailed than four-digit Dictionary of Occupational Titles or Census codes. For example, secretaries are divided into five occupation classes, depending on their responsibilities, and distinguished from other clerical occupations such as stenographers (three classes), typists (three classes), and file clerks (four classes). This level of detail provides strong control for human capital as productively used. (Groshen [1988b] tests this assertion.) For brevity in the discussion that follows, the term occupation will be used instead of AWS job classification, the more accurate term.

In total, the particular survey analyzed below covers 88 occupations and 241 establishments in 42 two-digit SIC categories. Confidentiality restrictions prevent the Bureau of Labor Statistics from releasing the identity of the MSA or the exact years covered. The MSA is described as located in the northeast region of the country and not widely dispersed geographically. The six consecutive years fall between 1974 and 1981.[6]

Table 1 presents a summary of characteristics of the sample. Almost half (108) of the establishments are covered for the full six years; the remainder are fairly evenly split between those present for the first three years and the last three years, except for the few (7 percent) with missing data for one or more years. Thus, the data cover 1,008 establishment-years. In any year, well over half of the establishments are among those covered for the full six years. Approximately 17,000 individuals are surveyed per year, for a grand total of 101,990 observations.

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

Because the AWS occupations are found in many different industries and firms, the labor markets for such occupations may be more competitive than the markets for more industry-specific or firm-specific occupations. Workers can be expected to be more mobile when their skills are readily transferable among many different employers. Thus, we would expect the wages of workers in AWS occupations to be more standard across employers than would the wages of workers in less common occupations.

However, because they are common to most firms, AWS occupations generally work outside the major productive activities of the establishments surveyed and capture a relatively small proportion of the employees in most establishments. There are two alternatives to this approach. The first, analysis of industry-specific surveys that include the occupations most prevalent in each industry, is taken by Groshen (1988b). The second solution is to contract job classifications into broad occupational categories and survey all occupations and industries, as is done in household surveys. The analysis presented here includes a comparison of the results from the AWS to those from industry surveys and from the Current Population Survey.

III. The Size and Stability of Employer Wage Differences

A. Technique

Of particular interest in the study of interemployer wage differences is a measure of their *importance,* that is, the relative contribution of employer wage differences to total wage variation. This section partitions the variance of wages into the portions associated with particular effects using analysis of variance (ANOVA) techniques.

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

At any time, wages are hypothesized to depend on an individual's occupation, employer, the interaction between employer and occupation, and an individual component. If virtually all productive differences in human capital and working conditions are between, not within, narrowly defined occupations, then occupation dummies capture all significant differences in human capital and working conditions among jobs. Groshen (1988b) examines this issue and finds that the standard human capital variables (age, education, and race) add little explanatory power to regressions with three-digit occupational dummies in the Current Population Survey. Given the detail of the occupation distinctions in these surveys, the human capital variables *can* be expected to explain even less of the remaining variation in these data. In order to control as fully as possible for differences in worker quality, the actual estimation includes dummies for sex and incentive pay along with occupation. For ease of exposition, this set of variables is referred to simply as "occupation."

The test for the importance of employer characteristics is to measure the size and significance of employer variables included in a wage equation with human capital variables. The first set of variables are establishment dummy variables, to capture the average deviation of employees from their occupation means across all occupations. This effect, the fixed effect of employer on wages, is the main focus of this analysis.

Second, variations in employer differentials among occupations are captured by including variables for the interaction of occupation and establishment, which estimates an additional wage differential for each occupation in each plant. In this paper, this will be called an employee's "job-cell."

The equation estimated is as follows:

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

(1) $$W_{ijk} = \mu + X_i\alpha + Y_j\beta + X_iY_j\tau + \epsilon_{ijk},$$

where w_{ijk} = ln(wage) of employee k in occupation i at employer j,
μ = mean wage for the population,
X_i = vector of occupation dummy variables,
a = vector of occupation wage differentials,
Y_j = vector of establishment dummy variables,
β = vector of employer wage differentials,
X_iY_j = vector of job-cell dummy variables,
τ = vector of wage differentials for job-cells, and
ϵ_{ijk} = randomly distributed error term.

Since all of the independent variables are dichotomous, equation (1) can be rewritten, and wages may be understood, as the *sum* of a series of differentials:

(2) $$w_{ijk} = \mu + \alpha_i + \beta_j + \tau_{ij} + \epsilon_{ijk},$$

where α_i, β_j, and τ_{ij} are the i^{th}, j^{th}, and ij^{th} elements of the a, β, and τ vectors, respectively, and μ is the overall mean wage. Over time, any of these four components may change, introducing coefficients on their interactions with year. These year-interaction coefficients capture trends or temporary deviations from average relative position aver the *six* years and may be estimated in an expanded version of equation (2):

(3) $$w^t_{ijk} = \mu + \alpha_i + \alpha^t_i + \beta_j + \beta^t_j + \tau_{ij} + \tau^t_{ij} + \epsilon^t_{ijk}.$$

The differentials can be understood as follows:

1) Occupation differential (α_i) is an occupation's average deviation from mean wages, across all establishments. Presumably, these premia reflect productivity and compensating differences among occupations.

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

2) Occupation-year differential (α^t_i) is an occupation's average deviation from its own mean wage in a particular year, across all establishments. These movements reflect responses to temporary labor supply shocks or adjustments toward new long-run positions.

3) Establishment differential (β_j) is the employees' average deviation from occupation mean in an establishment, across all occupations. Thus, these encompass many differentials proposed in earlier research: size of employer, industry, percentage female, union, etc.

4) Establishment-year differential (β^t_j) is the employees' average deviation from establishment mean in a particular year, across all occupations. These movements reflect responses to temporary shocks or adjustments toward new long-run positions.

5) Job-cell (interaction) differential (τ_{ij}) is paid to a particular job-cell above the occupation and establishment differentials. High variance in this term indicates significantly different internal wage structures among employers.

6) Job-cell-year differential (τ^t_{ij}) is the job-cell deviation from mean in a particular year. High variance in this term indicates instability in the internal wage structures of employers.

7) Within job-cell (individual) differential (ϵ^t_{ijk}) is an individual or residual deviation from the mean for an occupation in an establishment in a year, presumably the result of individual productivity differences or differing compensation strategies on the part of employers (for example, incentive versus day rates). The more that wages are tied to individuals or to short-run performance rather than to jobs, the larger is this component.

Note that equations (1) and (2) express the same model in different notation. Equation (3) estimates the same model as in equations (1) and (2), but is fully interacted with time. If the differentials in equation (3) are mutually independent (this issue will be considered below), the total variance of wages may be partitioned as follows:

$$(4) \qquad \sigma^2_w = \sigma^2_\alpha + \sigma^2_{\alpha t} + \sigma^2_\beta + \sigma^2_{\beta t} + \sigma^2_\tau + \sigma^2_{\tau t} + \sigma^2_{\epsilon t}.$$

The size of each variance component estimate indicates its relative economic importance. And, the variation associated with interactions between a component and year measures the stability of wage differentials associated

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

with the component over time. Our interest is the economic and statistical significance of the differentials as groups, summarized by the relative size of the variance components and their interactions, as follows:

1) σ^2_a and σ^2_{at} measure the *importance* and stability of external occupational labor markets, respectively;

2) σ^2_β and $\sigma^2_{\beta t}$ measure the *importance* and stability of employer wage differentials in wage determination, respectively;

3) σ^2_τ and $\sigma^2_{\tau t}$ measure the importance and stability of independent internal labor markets, respectively; and

4) $\sigma^2_{\epsilon t}$ measures the importance and stability of individual differences within job-cell.

The essential complication to the discussion above is that variance-component decomposition as shown in equation (4) is not straightforward when data are unbalanced. An unbalanced design produces multicollinearity between the vectors of dummy variables (X_i and Y_j) in equation (1), which prevents a simple separation of the impacts of X and Y. If an establishment employs a relatively large number of workers in skilled occupations, we cannot distinguish whether a differential paid to those workers is due to their employer or to their occupations.[7]

Thus, the technique applied is a decomposition of the sum of squares of wages, rather than an explicit estimation of variance components.[8] This method provides a measure of the ambiguity arising from design imbalance and does not require the imposition of structure on estimated differentials.

The summary of the technique provided in table 2 shows how a series of ordinary least squares (OLS) regressions is used to make the jump from equa-

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

tion (3) to equation (4). Wages are regressed successively on different sets of regressors. Changes in the coefficient of determination (that is, the *sum* of squares explained as a proportion of total) are used to partition the *sum* of squares of wages into components corresponding to those in equation (4).[9] Use of the R^2 standardizes a^2_w to a value of one.

First, in the pooled sample, log wages are regressed separately on vectors of occupation and establishment dummies and then on both *sets* of dummies together (called the full main-effects model). The marginal contribution of each set of dummies to the full main-effects model (over the equation with the other one alone) measures the portion of wage variation associated unambiguously with that factor. These correspond to minimum estimates of the relative size of the variance contributed by occupation and differentials, or a^2_α and a^2_β. The difference between the R^2 of each in the equation alone and their marginal contribution to the full main-effects equation is a measure of their joint (collinear, or ambiguous) explanatory power. To identify the industry effect, industry dummies are substituted for establishment dummies.

Next, the exercise is repeated with interactions between the main effects and year, in order to estimate the relative size of σ^2_{at} and $\sigma^2_{\beta t}$, which indicate the stability of the main-effect estimates. The contribution of all other interaction differentials, including job-cell (a^2_τ) and job-cell-year differentials (a^2_{7t}), is the difference between the explanatory power of a regression on job-cell-year dummies and that of the full (time-interacted) main-effects model. The individual contribution ($a^2_{\epsilon t}$) is the variation unexplained by job-cell-year dummies.

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

B. ANOVA of the Area Wage Survey

Table 3 presents the ANOVA of wage data from the area wage survey. The first column reports the degrees of freedom for each source of variation.[10] The second column reports the percentage sum of squares, or increment to <, captured by each source. The total sum of squares reported excludes the effect of annual means, which were extracted prior to the analysis presented. The third column records F-statistics where appropriate.

The top six rows summarize the impact of the main effects: job classification, sex, incentive and establishment.[11] Together, these factors account for 90 percent of the observed variation in wages. The joint contribution of the main effects dominates, claiming 51 percent of total variation. This reflects an uneven distribution, or incomplete overlap, of occupations among establishments in the sample. The marginal contributions of establishment over occupation, and vice versa, are 19.3 percent and 19.5 percent, respectively: about equal, and both highly significant statistically. Each explains somewhere between 19 and 71 percent of total wage variation (71 percent is the marginal contribution plus the joint portion of variation).

The fixed establishment component of variation can be divided into the portions between industries and within industry. Between-industry variation is 11.4 percent of total variation (almost 60 percent of the marginal establishment total), leaving 7.9 percent for within-industry variation. Both portions have significant F-statistics. So, while industry captures a large part of the differences between establishment, it does not capture it all.

These results indicate that large establishment differentials exist within MSAs. The estimated establishment differentials have a large range:

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

from a minimum of -.81 to a maximum of +.56, compared to the mean. In fact, we cannot reject the possibility that employer differentials a .as important as occupation, sex, and incentive pay in the determination of wages.

The importance of the interactions with time and between occupation and establishment are examined in depth below. The final category is individual variation, which accounts for only 3 percent of total wage variation. This suggests that individuals in the same job-cell are paid very similarly.[12]

C. The Uniformity of Establishment Differentials Across Occupational Groups

The tenth row of table 3, "all other interactions," measures the contribution of all interactions not explicitly listed in the rows above. These interactions include job-cell and job-cell-year interactions (which measure a^2_7 and a^2_{7t}): differences in age-earnings profiles, in the relative treatment of job-cells by establishment, and changes in these over time. This group of interactions is significant as a whole, but accounts for just 6.3 percent of total variation. That is, the most conservative estimate of the contribution of employer main effects —19 percent—is three times as large as the interaction contribution. The size of this term suggests that relative occupational wage structures are probably fairly similar among these establishments.

Another way of examining the consistency of establishment differentials across occupational groups is to obtain and compare independent estimates for the four general occupational groups in the sample. Correlations of the employer wage differ —— across groups are shown in table 4.

The upper panel lists correlations across groups when industry effects are included in establishment effects. For instance, the correlation between the establishment differentials of office workers and those of maintenance,

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

toolroom and powerplant workers is .635. The correlations are similar in magnitude to those obtained in Leonard (1987) and Groshen and Krueger (1989). Rank order correlations (listed below the standard Pearson correlations) do not differ substantially.

The lower panel shows the cross-occupational consistency of establishment effects within industry. Again, the correlations are generally quite high. In fact, the correlations involving professional and technical workers rise after controlling for industry. The smallest correlation (.306) occurs between office occupations and material movement and custodial workers. Apparently, interindustry differentials account for the bulk of the consistency in interestablishment differentials between these two groups.

In general, though, these results suggest that establishment differentials have consistent size and rank across occupations, even within industry.

D. The Stability of Establishment Wage Differentials

The pattern of establishment and occupation wage levels in this survey remains unchanged over six years. This *can* be inferred from rows 7 through 9 of table 3, which suggest that occupation and establishment differentials are remarkably stable: occupation and establishment interactions with year contribute a total of less than 1 percent of observed variation. Employer differentials are only slightly less stable than occupation differentials.

Another demonstration of the stability of establishment differentials is the lack of decay in year-to-year correlations as the gap between observations lengthens. Table 5 presents correlation coefficients (both Spearman and Pearson) of estimated establishment differentials across time. The correlation coefficients of estimated differentials for the *same* establishments in different years are strikingly high, starting at .99 for one-year differences

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

and barely decaying to .97 for estimates six years apart. The picture for rank-order correlations is much the same: coefficients decline only to .95 after six years.[13]

The lower panel of table 5 shows the persistence of within-industry establishment differentials. Although somewhat lower than the persistence of differentials that include industry effects, the correlations are still remarkably high: they decline only to .894 (.856 in rank order) over the course of six years.

So, not only are employer differentials stable in size over time, but the relative rank of employers by size of differential is also stationary for periods as long as six years. Furthermore, the lack of any rapid decay over the period suggests that the patterns are probably stable for much longer than the six years included in the survey.

E. Conversion into Standard Deviations

Table 3 partitions the *sums* of squares, but does not indicate estimated variances for the components of interest. Table 6 presents results of multiplying the percentage of the sum of squares due to each factor by the total variance of the sample, and then taking the square root to generate the suggested standard deviation. In order to stack the deck against the investigated effect, the joint effects from table 3 are allocated completely to occupation.

The results can be converted to standard deviations in two ways. First we see the entire establishment effect, including the industry effects. This generates a standard deviation of .18, which we can interpret as a percentage of the mean because wages were estimated in log form. We can also *extract* two-digit industry effects from the estimated establishment effects.

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

This leaves intra-industry variation with a standard deviation of 12 percent, strikingly similar to the estimate of 11 percent in the industry surveys in Groshen (1988b). The similarity of these results, despite the very different sources, lends confidence to the findings.

Huw big are these numbers in practical terms? The experiment that this research tries to simulate is the random transfer of a worker in one establishment to a job in the *same* occupation at another establishment. What is the expected wage change from *such* a switch?[14]

Converting the suggested *standard* deviations in table 6 to expected wage changes, a random switch in establishment within industry (within job classification, sex, city, and incentive class) yields an expected 12 percent change (in absolute value) in wages; a switch that might be between industries is expected to generate a 19 percent wage change. These differences are comparable to average wage differences between union and nonunion employers, and correspond to differences of $2,100 and $3,300 per year, respectively, of the average wage of $17,000 earned by a blue-collar production worker in manufacturing in 1984. Switching employers within industry results in a very large expected income change, as large as that from a switch in occupation within industry. In addition to the stability they show, the *sheer* size of these differentials makes it unlikely that they are caused by random variations.

F. Employer Differentials and Wage Variation in the Current Population Survey

A large portion of current research in labor economics is based on log wage regressions of Current Population Survey (CPS) data, but at least half of

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

the wage variation in the CPS remains unexplained after inclusion of traditional measures of human capital. What portion of that "unexplained" variation is actually due to employer differentials?

Appendix B compares variance components estimates for the six-industry Industry Wage Survey (IWS) average in Groshen (1988b), for the AWS, and for the May 1977 CPS. The IWS estimates are the simple means from ANOVA of the wages of production workers in six manufacturing industries. The AWS estimates are repeated from table 6, except that the effects of all interactions with time have been removed.

Since these three data sources are quite different, adjustments for the differences are necessarily speculative. The conclusion reached is that, compared to total wage variation in the CPS, estimated variation due to establishment differentials is large, even by conservative measures.

IV. Establishment Size, Growth. and Shrinkage Differentials

The employer wage differentials estimated above are presumably linked to characteristics of the employers, some of which have been identified, such as size of firm and size of establishment (Brawn and Medoff [1987]). This section investigates the link between wages and another characteristic of establishment--growth or shrinkage of employment.

In these data, growth and shrinkage dummy variables *can* be created from changes over time in size class. Since Leonard (1989) finds that the size of establishment is surprisingly volatile, the first attempt to measure the influence of size change on wages uses net change in the size of employ-

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

ment at an establishment to measure growth and shrinkage. Dummies for growth and shrinkage were entered separately in order to allow for lack of symmetry in lags or for stickiness in either direction.

The upper panel of table 7 compares the contribution to explanatory power of size and the size change (row 3 of the table) to that of establishment dummies (row 2), controlling for occupation and industry (row 1). The purpose is to measure how much of employer variation within industry can be linked to size and size change. The results indicate that establishment size alone and dummy variables for establishment growth and shrinkage account for more than 19 percent of within-industry wage variation by employer in the AWS. Only 3 percent of this is contributed by the growth and shrinkage variables.

The lower panel of table 7 presents the coefficient estimates for the regression equation in row 3 in the upper panel. Except for the smallest size class, wages increase monotonically with size, and we estimate a negative differential for growth and a negligible one for shrinkage.

Table 8 presents the results of four other attempts to link estimated establishment differentials and changes in estimated differentials to growth or shrinkage of the establishment. The question is whether size change leads to greater or smaller wage changes than would be expected just from the adjustment to wages of the new size class.

If *growth* or shrinkage is exogenously determined and information is costly, then an employer's *growth* may raise its efficient wage under the turnover version of the efficiency wage hypothesis (see Salop [1979]). The wage increase is profit-maximizing because, during growth, the employer needs to attract or retain a higher proportion of workers than it does in a steady

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

state. Similarly, an employer that needs to shrink its work force may allow relative wages to fall below previous levels. Attraction of new workers is unnecessary and quits are perhaps desirable.

A second explanation for the *same* association comes from the bargaining model. Suppose that establishment *growth* resulted from success--that is, high profits--and shrinkage from low profits. Then, growth would indicate the presence of high wages because large rents were available for distribution. By the *same* logic, shrinkage would indicate lower wages.

However, if growth is endogenous, the zero-sum aspect of bargaining raises the possibility of the opposite relationship. If profits captured by workers would otherwise be used for expansion, then high-growth companies could be those with low wages. And shrinkers could be doing so because of their high wages. This is the *same* prediction and causality generated by the simple competitive model in the short run. Low wages lead to higher profits and, therefore, *growth*, unless the low wages induce quits, and thus, shrinkage; high wages should erode profits and cause shrinkage. Included here is the observation that since most hires are at the bottom of pay ranges, a hiring surge could appear to lower wages by lowering average tenure in a plant.

To summarize, the turnover version of the efficiency wage hypothesis predicts a positive relationship between *growth* and wages. The bargaining model is ambiguous, depending on the exogeneity of growth, and the simple competitive model predicts a negative relationship, or none at all.

The first two columns of table 8 present regression coefficients for the effect of establishment growth and shrinkage on estimated establishment differentials, controlling for industry and size. The effect of shrinkage may be negative, occurring before the shrinkage takes place. The effect of growth

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

is also negative, but relative to the wages of establishments in the new size class, not the old one. This suggests that wage changes may lag behind growth, but precede shrinkage. That is, wages may be sticky upwards during size change. Since the coefficient on growth is small and insignificant relative to that on past size, and the coefficient on shrinkage is small and insignificant relative to that on current size, the movement in wages is apparently not greater than that associated with a change of size category.

However, the third column of table 8 diminishes confidence in the last point. In order to allow for more complete adjustment and to increase the signal-to-noise ratio, this column presents regressions of net changes in estimated differentials on net changes in size. Neither growth nor shrinkage has a large or significant impact on change in differentials. The sign of the coefficient on shrinkage switches to positive but is small. Growth is estimated to reduce wages by 1 percent (with no controls for size), but the estimate is not significantly different from zero.

These data do not conclusively support any of the three hypotheses above. The first two columns suggest that wages are sticky upwards. If anything, wages are apparently lower for firms that grow, but shrinkage has little or no effect. And, neither result is stable under alternative formulations (that is, relative to wages of employers of the same size).

Thus, although size changes affect wages because wages increase with size, neither growth nor shrinkage appears to have a simple, consistent effect on wages, holding size constant. The data reject the efficiency wage and the exogenous-growth bargaining predictions of a positive relationship between growth and wages. The correlation between wages and growth, if there is one,

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

appears to be negative. It is even less likely that shrinkage is correlated with wages; but if so, shrinkage is also associated with (slightly) lower wages.

V. Conclusion

A. Summary of Findings

The conclusions of this analysis are as follows:

(1) Twenty to 70 percent of wage variance within this MSA is due to employer-based differences both between and within industry. The most conservative estimate of the standard deviation due to employer differentials within industry is 12 percent. Combined with industry effects, this generates a standard deviation of approximately 18 percent: a major portion of the 50 percent total standard deviation of wages.

(2) Establishment wage differences and rankings (even within industry) are virtually stationary for periods at least as long as six years, and probably for longer.

(3) While establishment size can account for much of measured employer wage effects within industry, establishment *growth* and shrinkage do not have a simple, consistent relationship with employer wage levels or wage *changes.*

Thus, even across occupations as diverse as those in the area wage survey, employer differentials are applied relatively uniformly. Compared to occupational means, employers tend to compensate janitors as well (or as poorly) as they do industrial nurses, computer programmers, millwrights, and stenographers. Furthermore, employers are also very consistent in their patterns aver time.

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

Occupation (including sex and incentive) and employer differentials are clearly extremely important in wage determination. These factors, when well identified, as in these surveys, can explain more than 95 percent of wage variation. Thus, other characteristics of the individual (for example, tenure, marital status, or race) must operate through job classification or through employer in order for them to have a large effect on wages. Otherwise, they are not highly influential in the determination of wages.

In short, since a large improvement in earnings can be attained only through a promotion or a change of employer, barriers to entry into highly remunerative occupations or establishments *can* have a devastating impact on the earnings of otherwise-qualified workers.

B. Implication for Sources of Establishment Wage Differentials

These results cast more light onto the nature of wage differences among employers and onto the plausibility of proposed sources of wage variation by employer. Of the five sources of employer wage differentials that have been modeled (sorting by worker quality, compensating differentials, random variations, efficiency wages, and insider bargaining), evidence in previous studies renders the first two possibilities unlikely.

The evidence presented above rejects the third possibility, random variations, as the *source* of employer differentials. The strong stability of establishment differentials over time provides compelling evidence against the hypothesis that establishment differentials are temporary fluctuations. If the differentials are random but not temporary, then they are extremely costly for high-wage employers, which suggests that labor-market information must be

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

even more costly. But, the results of this survey and many other private substitutes are available to firms on a fairly timely basis at no cost (or at the cost of participation).

Furthermore, the extent to which the differentials persistently (since at least the 1940s) depend on easily identified establishment characteristics, such as industry and size of establishment, makes the randrnn variations hypothesis unlikely. For instance, it is implausible that personnel officers of large firms have been consistently wrong, all mistakenly setting their wages too high for 40 years. Thus, the randrnn variation theory of establishment differentials may be ruled out.

The finding of substantial wage differences among employers within a single city also argues against the possibility that regionwide compensating differentials for cost of living are the main source of employer differentials, although urban wage gradients within the city are still a possibility.

Process of elimination also suggests the need for serious consideration of efficiency wage and rent-sharing (insider/outsider) models. This paper identifies several key characteristics of interemployer wage differentials that need to be present in any version of these models invoked.

First, employer wage differentials are found among white-collar workers, as well as blue-collar workers, in a nationally representative set of industries. The pervasiveness of these differentials argues for explanations that apply across-the-board to all occupations in an establishment, and to the establishments in most industries. Thus, occupation-specific difficulties in monitoring are not a likely source, because the occupations surveyed here are very diverse. Also unlikely are explanations that appeal to the characteristics of a single industry.

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

Second, although wages and size of establishment have a strong positive correlation, plant size change has no simple, consistent relationship with wage level. Thus, the versions of efficiency wage and rent-sharing models based on growth or shrinkage of establishment are unlikely sources of interemployer wage differences.

Third, since employer differentials are quite persistent on an annual basis, while annual profit rates of U.S. companies are notoriously volatile, if these differentials are r e n t . , they presumably reflect long-run, not short-run, rents.

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

References

Akerlof, George, "Labor Markets as Partial Gift Exchange," Quarterly Journal of Economics, vol. 97, November 1982, pp. 543-569.

Aoki, Masahiko, The Cooperative Game Theory of the Firm, New York: Oxford University Press, 1984.

Brown, Charles, "Equalizing Differences in the Labor Market," Quarterly Journal of Economics, Vol. 94, No. 1, February 1980.

_____ and James L. Medoff, "The Employer Size Wage Effect," unpublished paper, September 1987.

Brown, William, John Hayles, Barry Hughes and Lyndon Rowe, "Production and Labor Markets in Wage Determination: Same Australian Evidence," British Journal of Industrial Relations, Vol. 22, No. 2, July 1984.

Bulow, Jeremy I., and Lawrence H. Summers, "A Theory of Dual Labor Markets With Application to Industrial Policy, Discrimination, and Keynesian Unemployment," Journal of Labor Economics, Vol. 4, No. 3, Part 1, July 1986, pp. 376-414.

Dickens, William, "Wages, Employment and the Threat of Collective Action by Worker," National Bureau of Economic Research Working Paper No. 1856, March 1986.

_____ and Lawrence F. Katz, "Industry Wage Patterns and Theories of Wage Determination," National Bureau of Economic Research Working Paper No. 2271, July 1987.

Doeringer, Peter, and Michael Piore, Internal Labor Markets and Manpower Analysis, Lexington, Mass.: D.C. Heath and Co., 1971.

Dunlop, John T., "The Task of Contemporary Wage Theory," in G. Taylor and F. Pierson, eds., New Concepts in Wage Determination, New York, 1957.

Eberts, Randall W., "An Empirical Investigation of Intraurban Wage Gradients," Journal of Urban Economics, Vol. 10, 1981, pp. 50-60.

Gibbons, Robert, and Lawrence F. Katz, "Learning, Mobility, and Inter-Industry Wage Differentials," Massachusetts Institute of Technology, unpublished paper, December 1987.

Groshen, Erica L., "Sources of Wage Dispersion: Does It Matter Where You Work?" Ph.D. Dissertation, Harvard University, 1986.

______, "Why Do Wages Vary Among Employers?" Federal Reserve Bank of Cleveland Economic Review, First Quarter 1988a.

______, "Sources of Wage Dispersion: The Contribution of Interemployer Wage Differentials Within Industry," Federal Reserve Bank of Cleveland Working Paper No. 8802, 1988b.

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

_____ and Alan B. Krueger, "The Structure of Supervision and Pay in Hospitals," Federal Reserve Bank of Cleveland Working Paper No. 8907, June 1989.

Henderson, C.R., "Estimation of Variance and Covariance Components," Biometrics, June 1953.

Hocking, Ronald R., O.P. Hackney, and F.M. Speed, "The Analysis of Linear Models with Unbalanced Data," in H.A. David, ed., Contributions to Survey Sampling and Applied Statistics: Papers in Honor of H. O. Hartley, New York: Academic Press, 1978.

Leonard, Jonathan S., "On the Size Distribution of Employment and Establishments," Quarterly Journal of Economics, 1989.

______, "Carrots and Sticks: Pay, Supervision, and Turnover," Journal of Labor Economics, Vol. 5, No. 4, October 1987, pp. s136-s152.

Mackay, Donald I., David Boddy, John Brack, John A. Diack and Norman Jones, Labour Markets Under Different Employment Conditions, London: George Allen & Unwin Ltd., 1971.

Nolan, Peter and William Brown, "Competition and Work Place Wage Determination," Oxford Bulletin of Economics and Statistics, Vol. 45, No. 3, August 1983.

Rothschild, Michael, and Joseph E. Stiglitz, "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," Quarterly Journal of Economics, Vol. 90, No. 4, November 1976.

Salop, Steven, "A Model of the Natural Rate of Unemployment," American Economic Review, Vol. 69, No. 1, March 1979.

Searle, S.R., Linear Models, New York: Wiley & Sons, 1971.

Segal, Martin, "Post-Institutionalism in Labor Economics: The Forties and Fifties Revisited," Industrial and Labor Relations Review, Vol. 39, April 1986, pp. 388-403.

Smith, Robert, "Compensating Wage Differentials and Public Policy: A Review," Industrial and Labor Relations Review, Vol. 32, No. 3, April 1979.

Stigler, George J., "Information in the Labor Market," Journal of Political Economy, Vol. 70, No. 5, Part 2, October 1962.

Stiglitz, Joseph E., "Theories of Wage Rigidities," National Bureau of Economic Research Working Paper, 1984.

_____, "On Search and Equilibrium Price Distributions," in Economics of Human Welfare, Michael J. Posken, ed., New York: Academic Press, Inc., 1979.

Topel, Robert, "Equilibrium Earnings, Turnover and Unemployment," Journal of Labor Economics, Vol. 2, October 1984.

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

Weitzman, Martin, The Share Economy: Conquering Stagflation, Cambridge, Mass.: Harvard University Press, 1986.

Yellen, Janet, "Efficiency Wage Models of Unemployment," American Economic Review, Vol. 74, No. 2, May 1984.

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

Footnotes

1. Groshen (1988a) reviews the empirical and theoretical literature examining employer differentials.

2. Exceptions to this generalization are a group of studies by economists in the 1940s and 1950s summarized in Segal (1986). Groshen (1988b) provides recent evidence of large establishment wage differentials among production workers in six manufacturing industries, using national industry wage surveys.

3. Groshen (1988b) finds it unlikely that intra-industry employer variations are due to sorting by tenure, experience, education, or for variations in unmeasured worker ability correlated with these measures of human capital. Dickens and Katz (1987) find that interindustry differentials cannot be explained by the three measures of human capital. And, Gibbons and Katz (1987) conclude that interindustry wage differences are not associated with unmeasured differences in productive abilities.

4. Attempts to identify the working conditions for which interindustry wage variations compensate have been notably unsuccessful, as have attempts to identify compensating variations in general (Brown [1980] and Smith [1979]).

5. However, urban wage gradients within the city are still possible (Eberts [1981]).

6. These years were characterized by historically high inflation rates, which might be expected to result in more random behavior because of more costly information, and in more real downward wage flexibility on the part of employers.

7. Techniques for estimation of variance components of a model of unbalanced design are detailed in Searle (1971) and Henderson (1953). Restricted maximum likelihood (RML) techniques are introduced in Hoching, Hackney and Sped (1978). RML provides simple estimates of variance components and their standard errors at the *expense* of imposing a rigid structure on the distribution of level effects and errors. Because the appropriateness of the structure imposed may vary among industries, and because the purpose of this study is to investigate the characteristics of establishment differentials, a nonparametric method was preferred for this analysis. Groshen (1986) provides a camplete discussion and examples of the application of alternative ANOVA techniques to similar data.

8. The technique used here avoids the essence of ANOVA's difficulty with unbalanced data. A variance is a *sum* of squared deviations divided by the appropriate number of observations or degrees of freedom. In data with an unbalanced design, the correct number of degrees of freedom is unknown, so variance estimates must rely on estimates of the correct degrees of freedom. Such estimates require the imposition of structure on the data.

9. The following work concentrates attention on proportions of variance rather than on F-statistics for two reasons. First, because of the large sample sizes, all of the F-statistics are strongly significant (the critical value is 1 in most cases), even if the economic significance is slight. Second, establishment identity is presumably an inefficient measure of the economically

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

relevant differences between establishments. By construction, it captures all differences and thus identifies the maximum amount of variation that understanding of employer wage policy could explain.

Hwever, as a measure of the source of employer differences, establishment may be finer than necessary. If so, the F-statistic can mislead because it averages out the impact of all estimated levels. While the additional variation explained by unnecessary levels is negligible, the number of degrees of freedom used can be high, reducing the F-statistic. The inclusion of irrelevant levels washes out the significance of the relevant ones.

The F-statistic of a factor X is defined as follows:

$$F_x = [(RRSS-URSS)/k]/[URSS/(n-k)],$$

where RRSS = restricted residual *sum* of *squares*, URSS = unrestricted residual *sum* of ***squares***, k = number of restrictions or levels in parameter X, and n = degrees of freedom in unrestricted equation (that is, number of observations minus degrees of freedom used by other regressors).

If k is the number of correctly specified levels of the factor X, then let 6 = measure of irrelevant fineness in another measure, say Y. That is, suppose instead of using k levels, we use the δk levels of Y, where $\delta>1$. Then, as long as the levels of X are a linear combination of the levels of Y, and n is large relative to δk, the URSS of the equation will be almost the same, the RRSS will be the same, so the F-statistic of the inefficient parameter Y is as follows:

$$F_y \cong [(RRSS-URSS)/\delta k]/[URSS/(n-\delta k)].$$

And, the ratio of F_y to F_x (for n large relative to k) is

$$F_y/F_x \cong (n-\delta k)k/[(n-k)\delta] = [(n/\delta)-k]/(n-k) \cong 1/\delta.$$

The maximum of the ratio is one (where X=Y, so $\delta=1$); otherwise it decreases monotonically with increasing 6, and approaches $1/\delta$ for n large and k small. So the size of the F-statistic depends not only on the economic relevance of the parameter measured, but also on the inefficiency with which it is measured. Since the purpose of this work is to identify the potential explanatory p e r of variables based on establishment, I focus primarily on the percentage *sum* of squares explained by factors rather than through F-statistics.

10. The number of degrees of freedom is determined by the number of dummy variables used in the regressions. For example, in the case of establishments, the number of degrees of freedom is the number of establishments minus one.

11. The incentive dummy equals one when the worker in question has an incentive component to his or her earnings. These incentives may be in the form of individual or *group* piece rates, individual or *group* bonuses, or commissions.

12. This is the result for industries with a low proportion of incentive-based compensation in Groshen (1988b).

13. These are quite similar to the results obtained by Mackay, et al. (1971) and Nolan and Brown (1983) in England.

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

14. This question asks for the expected absolute value of the difference between two identically distributed random variables. Assuming a normal distribution of differentials, the question reduces as follows:

$$E[\Delta d] = E[|d_1 - d_2|] = 2^{1/2}(\phi[0]/\Phi[0])\sigma_d = 2^{1/2}(.4/.5)\sigma_d = 1.13\sigma_d,$$

where d = random differential, distributed $N(0,\sigma^2_d)$, and $\phi[0]$ and $\Phi[0]$ are the normal density and the cumulative normal density functions, evaluated at zero.

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

Appendix A

Occupations Surveyed in the Area Wage Survey

Office Occupations

Secretaries: Classes A, B, C, D, E, and Not Classifiable By Level
Stenographers: Senior, General, and Not Classifiable By Level
Transcribing-Machine Typists
Typists: Classes A, B, and Not Classifiable By Level
File Clerks: Classes A, B, C, and Not Classifiable By Level
Switchboard Operators
Switchboard Operator-Receptionists
Order Clerks: No Level Distinctions, Classes A, B, and Not Classifiable
Accounting Clerks: Classes A and B, and Not Classifiable By Level
Bookkeeping-Machine Operators: Classes A, B, and Not Classifiable By Level
Messengers
Billing-Machine Billers
Bookkeeping-Machine Billers
Machine Billers, Not Classifiable By Level
Payroll Clerks
Key Entry Operators: Classes A, B, and Not Classifiable By Level
Tabulating-Machine Operators: Classes A, B, and C

Professional and Technical Occupations

Computer Systems Analysts (Business): Classes A, B, C, and Not Classifiable
Computer Programmers (Business): Classes A, B, C, and Not Classifiable
Computer Operators: Classes A, B, C, and Not Classifiable By Level
Drafters: Classes A, B, C, and Not Classifiable By Level
Drafter-Tracers
Electronics Technicians: Classes A, B, C, and Not Classifiable By Level
Peripheral Equipment Operators
Computer Data Librarians
Registered Industrial Nurses

Maintenance, Toolroom and Powerplant Occupations

Main —— Carpenters
Maintenance Electricians
Maintenance Painters
Main —— Mechanics (Machinery)
Maintenance Mechanics (Motor Vehicles)
Maintenance Pipefitters
Maintenance Sheet-Metal Workers
Millwrights
Maintenance Trades Helpers
Machine-Tool Operators (Toolroom)
Tool and Die Makers
Stationary Engineers
Boiler Tenders

Material Movement and Custodial Occupations

Truckdrivers: Light Truck, Medium Truck Heavy Truck, Tractor-Trailer, and Not Classifiable by Category
Guards: No Level Distinction, Classes A, B, and Not Classifiable
Shippers
Receivers
Shippers and Receivers
Warehousemen
Order-Fillers
Shipping Packers
Material Handling Laborers
Forklift Operators
Power-Truck Operators (Other Than Forklift)
Janitors, Porters, and Cleaners
Watchmen

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

Appendix B

Decomposition of the Variance of Wages in Three Data Sets

This appendix presents variance components estimates for the six-industry Industry Wage Survey (IWS) average in Groshen (1988b), for the Area Occupational Wage Surveys (AWS), and for the May 1977 Current Population Survey (CPS). May 1977 was chosen as a year within the ranges of both the AWS and IWS. The sample includes all private-sector, full-time employees between the ages of 18 and 65 with reported average hourly earnings of more than $1.75 per hour.

The IWS estimates are the simple means from ANOVA of the wages of production workers in six manufacturing industries. The technique used in Groshen (1988b) is identical to that used here, except that all data are cross-sectional, and so differentials are estimated without explicit interactions with year. The AWS estimates are repeated from table 5, except that the effects of all interactions with time have been removed.

These three data sources are quite different, so adjustments for the differences are necessarily speculative. For instance, the standard deviation of wages in the AWS, .40, is double the mean for the six industry wage surveys (.20). As noted above, area wage surveys cover a broader mix of occupations, both blue-collar and white-collar. Moreover, area wage surveys include the effects of interindustry wage variation. The CPS includes all of the sources of variation already mentioned, in addition to the full range of occupations in the economy.

The first two rows of table B-1 present the least comparable numbers across the three surveys: standard deviation estimates for total dispersion

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

and those due to occupation, sex, region, and industry differentials. Reported AWS and IWS figures allocate the entire joint occupation establishment effect to occupation. In the IWS, the variance in the first row includes regional variation, but not interindustry variation. In the AWS, the variation in the first row includes interindustry variation, but no regional variation.

In the CPS, the first row captures both industry and regional sources of wage variation, in addition to occupation and sex. The level of detail of region, sex, and industry are roughly the same in the AWS and CPS, but CPS three-digit occupations lack the detail of the job classifications in the IWS and AWS. The CPS variation in the first row is the same as that of the AWS, despite the higher total variance in the CPS. This suggests that variation within the CPS occupational categories is greater than the variation between regions in the country. Lack of occupational specificity leaves more wage variation unexplained than the addition of regional controls can capture.

Another way to judge the impact of broad occupation data in the CPS is to note that in the plastics industry, contraction of the 42 BLS job classifications into 12 CPS occupational categories reduces the R^2 of the equation by one half, from 49 percent to 25 percent. In an ANOVA as shown, at least half of this difference—judging from the size of the contribution "joint" to occupation and establishment-might then be claimed by establishment differentials, raising the estimated employer effect in the CPS.

The second row shows the remaining variation for each sample. These are quite similar for the AWS and IWS: a standard deviation of about .16. The CPS, however, retains a standard deviation of .31, almost twice as high.

The next three rows present speculative estimates of the size of the within-industry establishment effect in the CPS, in order to provide bounds

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

for the probable contribution of establishment to CPS wage variation. The first method takes the point estimate of standard deviation from the IWS and AWS: .11. Although this is a large portion of the unexplained standard deviation of .31, the estimate is conservative for two reasons. First, CPS occupations are very broad. The large joint component of variation in the IWS and AWS would shrink with these broad occupations, increasing the size of the estimated establishment impact on variation. Second, the IWS and AWS oversample large establishments and omit the smallest ones. In these data, estimated establishment variance is highest among the smallest establishments. Thus, the CPS should provide more establishment diversity because it samples evenly from all sizes of employer.

The second estimate assigns the AWS establishment percentage of total wage variation to establishment in the CPS, and converts this to a standard deviation of .13. The result is very similar to the first estimate and has the same limitations.

The third method is less conservative and assigns to establishment the same percentage of remaining variation (after occupation, industry, etc.) as is found in the AWS. That converts to a standard deviation of .20.

In order to see if the limited number of occupations surveyed in the AWS accounted for these results, the last column of table 6 presents the *same* exercises on the subsample of CPS observations for workers in AWS occupations. (They totalled 24 percent of the CPS sample.) The variance of wages is lower in the subsample, but the entire decrease in variance is in the between-occupation portion of variance. This leaves the estimates of establishment effect virtually the *same*, increasing confidence in them.

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

But how much of the remaining variation is actually noise? The reasons CPS wage reports may have a larger noise-to-signal ratio than BLS wage surveys are as follows:

1) CPS average hourly earnings are somewhat imprecisely defined (they include earnings from overtime or shift premia or from second jobs);

2) CPS respondents' memories are probably subject to more error than are the establishment records used by the BLS;

3) CPS data-cleaning is far less thorough than BLS efforts; and

4) CPS occupations are subject to large reporting error.

So, the nonoccupation variation in the CPS is probably biased upwards. Thus, compared to total wage variation in the CPS, estimated variation due to establishment differentials is large, even by conservative measures.

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

Table 1

Characteristics of Area Wage Survey Sample

Mean Wage	$5.68
Variance of ln (Wage)[1]	.174
Standard Deviation of ln (Wage)[1]	.42
Number of Observations	101,990
Number of Occupations	88
Number of Establishments	241
Male	59.3%
Receive Incentive Pay	2.2%

Establishment Size	Percent of Observations
1-19	0.0%
20-49	0.5%
50-99	2.4%
100-249	9.3%
250-499	16.7%
500-999	13.4%
1,000-2,499	30.5%
2,500+	27.1%

Major Industry Group (1-Digit SIC)		Percent of Observations
2.	Nondurable Manufacturing	10.0%
3.	Durable Manufacturing	28.8%
4.	Transport. and Utilities	11.0%
5.	Wholesale and Retail Trade	17.3%
6.	Financial Services	12.8%
7.& 8.	Personal and Business Services	19.7%

Year of Observation	
1	17.3%
2	16.7%
3	16.5%
4	16.6%
5	16.5%
6	16.4%

Number of Years Observed	
1	3.0%
2	3.8%
3	46.4%
4	4.5%
5	1.5%
6	40.8%

[1]Net of annual effects.

Source: Tabulations from the BLS Area Wage Survey, unidentified area in the Northeast for six consecutive years between 1975 and 1982.

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

Table 2

Technique for Partitioning Sum of Squares in Unbalanced Data

Source of Variation	Percent of Total Sum of Squares[1]
1. Occupation, Sex, Incentive (controlling for estab.)	$R^2_C - R^2_B$
2. Joint Occupation and Establishment	$R^2_A + R^2_B - R^2_C$
3. Establishment and Industry (controlling for occup., etc.)	$R^2_C - R^2_A$
4. Industry (controlling for occupation, etc.)	$R^2_Ç - R^2_A$
5. Establishment Within Industry	$R^2_C - R^2_Ç$
6. Total Main Effects	R^2_C
7. Occupation, etc., -Year Interactions	$R^2_{CT} - R^2_{BT}$
8. Joint Occupation, etc., and Establishment	$R^2_{AT} + R^2_{BT} - R^2_{CT}$
9. Establishment Year-Interactions	$R^2_{CT} - R^2_{AT}$
10. All *Other* Interactions (controlling for main effects)	$R^2_D - R^2_C$
11. Total Between Job-Cell-Years	R^2_D
12. Individual	$100\% - R^2_D$
TOTAL	100%

[1]The subscripts on the coefficients of determination correspond to the regression models listed below. Occupation, sex, and incentive are listed as occupation, for ease of exposition.

A. $w_{ijk}{}^t = \mu + X_i\alpha + \epsilon_{ijk}{}^t$

AT. $w_{ijk}{}^t = \mu + X_i\alpha + X_i{}^t\alpha^t + \epsilon_{ijk}{}^t$

B. $w_{ijk}{}^t = \mu + Y_j\beta + \epsilon_{ijk}{}^t$

BT. $w_{ijk}{}^t = \mu + Y_j\beta + Y_j{}^t\beta^t + \epsilon_{ijk}{}^t$

C. $w_{ijk}{}^t = \mu + X_i\alpha + Y_j\beta + \epsilon_{ijk}{}^t$

Ç. $w_{ijk}{}^t = \mu + X_i\alpha + \underline{Y}_j\underline{\beta} + \epsilon_{ijk}{}^t$

CT. $w_{ijk}{}^t = \mu + X_i\alpha + X_i{}^t\alpha^t + Y_j\beta + Y_j{}^t\beta^t + \epsilon_{ijk}{}^t$

D. $w_{ijk}{}^t = \mu + X_i\alpha + X_i{}^t\alpha^t + Y_j\beta + Y_j{}^t\beta^t + X_iY_j\tau + X_i{}^tY_j{}^t\tau^t + \epsilon_{ijk}{}^t$

where $w_{ijk}{}^t$ = ln wage of individual k in occupation, establishment j, and year t
X_i = vector of occupation dummy variables for occupation i
Y_j = vector of establishment dummy variables for establishment j
$\underline{Y}_j$ = vector of industry dummy variables for industry j
X_iY_j = dummies for occupation i in establishment j, i.e., for job-cell ij,
a, β, $\underline{\beta}$, τ = vectors of estimated parameters, and
the superscript t denotes variables and parameters that vary over time.

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

Table 3

Analysis of Sources of Wage Variation Within an Area[1]

Source of Variation	Degrees of Freedom	Percent of Total Sum of Squares	F-Statistic[8]
1. Occupation, Sex and Incentive[2]	89	19.5%	2,168
2. Joint Occupation, etc., and Establishment	-	50.9	-
3. Establishment and Industry[3]	240	19.3	804
4. Industry[3]	41	11.4	1,558
5. Establishment Within Industry[4]	199	7.9	330
6. Total Main Effects	329	89.7	-
7. Occupation, etc. -Year Interactions[5]	436	0.3	8
8. Joint Occupation, etc. and Establishment	-	0.1	-
9. Establishment-Year Interactions[6]	767	0.5	7
10. All Other Interactions[7]	11,230	6.3	16
11. Total Between Job-Cell-Years	12,762	96.9	-
12. Individual	89,222	3.1	-
TOTAL	101,984	100.0%	-
Total Sum of Squares		15,934	

[1]All reported figures are net of main annual effects.
[2]Controlling for industry and establishment.
[3]Controlling for occupation, sex, and incentive.
[4]Controlling for occupation, sex, incentive, and industry.
[5]Controlling for main effects and establishment-year interactions.
[6]Controlling for main effects and occupation, sex, incentive-year interactions.
[7]Controlling for main effects and their interactions with year.
[8]All F-statistics are significant at well above the 1% level.

Source: Tabulations from BLS Area Wage Survey.

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

Table 4

Correlations of Estimated Establishment Wage Differentials
Over Four Occupational Groups[1]

A. Including Industry Effects[1]

	TYPE OF CORRELATION	Professional and Technical	Maintenance, Tool-Room and Powerplant	Material Movement and Custodial
Office	Pearson	.854	.635	.631
	Rank	.788	.670	.584
Professional and Technical	Pearson		.622	.503
	Rank		.636	.466
Maintenance, Toolroom, and Powerplant	Pearson			.773
	Rank			.787

B. Controlling for Industry Effects[1]

	TYPE OF CORRELATION	Professional and Technical	Maintenance, Tool-roam and Powerplant	Material Movement and Custodial
Office	Pearson	.886	.652	.306
	Rank	.892	.652	.134
Professional and Technical	Pearson		.799	.531
	Rank		.759	.486
Maintenance, Toolroom and Powerplant	Pearson			.732
	Rank			.611

[1]Results weighted by number of observations in establishment. Estimated establishment differentials are average differentials (taken from independent regressions for each occupational group) over period in which the establishment was observed.

Source: Tabulations from BLS Area Wage Survey.

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

Table 5

Correlations of Estimated Establishment Differentials Over Six Years

A. Including Industry Effects[1]

		TYPE OF CORRELATION	Year 2	3	4	5	6
	1	Pearson	.988	.988	.979	.969	.967
		Rank	.977	.978	.956	.937	.950
	2	Pearson	-	.989	.982	.978	.977
		Rank	-	.984	.968	.948	.962
Year	3	Pearson		-	.990	.981	.976
		Rank		-	.977	.952	.964
	4	Pearson			-	.991	.984
		Rank			-	.975	.971
	5	Pearson				-	.989
		Rank				-	.978

B. Controlling for Industry Effects[2]

		TYPE OF CORRELATION	Year 2	3	4	5	6
	1	Pearson	.975	.968	.909	.904	.894
		Rank	.970	.962	.891	.869	.856
	2	Pearson	-	.974	.925	.924	.906
		Rank	-	.969	.909	.897	.871
Year	3	Pearson		-	.950	.932	.909
		Rank		-	.916	.902	.877
	4	Pearson			-	.971	.949
		Rank			-	.967	.938
	5	Pearson				-	.959
		Rank				-	.969

[1]Results weighted by number of observations in establishment.
[2]Results weighted by number of observations in establishment. Industry-year effects are excluded. Establishments in industries with only one establishment are also omitted.

Source: Tabulations from BLS Area Wage Survey.

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

Table 6

Suggested Standard Deviations for Area Wage Survey

Source	Suggested Standard Deviation[1]
Occupation	.35
Establishment (Including Industry)	.18
Establishment (Within Industry)	.12
Interactions	.11
Individual	.07
TOTAL	.42

[1]Suggested standard deviation=[(category proportion of CSS)x(total variance)]$^{1/2}$. Joint contribution is allocated to occupation.

Source: Tabulations from BLS Area Wage Survey.

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

Table 7

Comparison of Regression on Establishment Dummies
with Regressions on Establishment Size in the Area Wage Survey

A. Comparison of Explanatory Power

Eq.	Independent Variables	R^2	ΔR^2 from Eq. 1
(1)	Occupation, Sex, Incentive and 2-Digit SIC	81.8	-
(2)	Occupation, etc. and Establishment Dummies	89.7	+7.9
(3)	Occupation, etc., SIC, Establishment Size Category and Net Size Change	83.3	+1.5
	RATIO OF EXPLANATORY POWER OF ESTABLISHMENT SIZE TO ESTABLISHMENT DUMMIES		.190

B. Coefficients from Regression of ln (Earnings) on Establishment Size

Variable	Coefficient or Number of Dummies	(Std. Error)
Occupation	87	
Male	0.047	(0.002)
Receive Incentive Pay	0.108	(0.004)
2-Digit SIC	41	
Establishment Size		
20-49	-0.172	(0.007)
50-99	-0.221	(0.003)
100-249	-0.193	(0.002)
250-499	-0.141	(0.002)
500-999	-0.093	(0.002)
1,000-2,499	-0.061	(0.002)
2,500+	-	
Net Shrinker	-0.013	(0.002)
Net Grower	-0.071	(0.002)
R^2	83.3	

Source: Tabulations from BLS Area Wage Survey.

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

Table 8

Effect of Establishment Size Change on Estimated Establishment Differentials in the Area Wage Survey

	Dependent Variable: Current Estimated Establishment Differential (1)	Dependent Variable: Current Estimated Establishment Differential (2)	Dependent Variable: Net Change in Estimated Establishment Differential Over Survey Period (3)
Coefficient on Establishment Shrinkage Dummy (std. error)[1]	-0.005 (0.024)	-0.052[2] (0.023)	0.004 (0.013)
Coefficient on Establishment Growth Dummy (std. error)[1]	-0.049[2] (0.024)	-0.011 (0.025)	-0.011 (0.017)
Other Controls	2-Digit SIC, Current Estab. Size	2-Digit SIC, Previous Estab. Size	Years Spanned[3]
R^2	0.712	0.711	0.513
F-Stat. for Size Changes	2.02	2.67	0.28
Sample Size	767	767	231
Weight	Number of observations in establishment		Average number of observations in establishment

[1]Growth and shrinkage are defined as positive or negative changes (respectively) in the establishment size category. For Equations 1 and 2, the change is from the last year to present. For Equation 3, it is net change over the survey period.

[2]Significant at the 5% level.

[3]Control for years *spanned* is necessary because the calculation and elimination of annual effects may introduce bias (due to sample variations) in year-to-year comparisons of wage effects.

Source: Tabulations from BLS Area Wage Survey.

http://clevelandfed.org/research/workpaper/index.cfm
Best available copy

Table B-1

Industry and Area Wage Survey Standard Deviation Components Compared to Current Population Survey Log Wage Variation

Source of Variation of Log Wage	Industry Wage Survey Mean Suggested Standard Deviation	Area Wage Survey Suggested Standard Deviation[2]	Current Population Survey May 1977[3] All Occup.	Current Population Survey May 1977[3] AWS Occup.[4]
Total Std. Dev.	.20	.40	.48	.42
Occupation, Sex, Region, and/or Industry[1]	.12	.36	.36	.25
Total Remaining	.16	.16	.31	.33
Establishment (known)	.11	.11	-	-
Establishment (Estimated)				
1) AWS & IWS Point Estimate	-	-	.11	.11
2) AWS % of Total	-	-	.13	.12
3) AWS % of Remaining	-	-	.20	.21
Occupation-Establishment Interaction	.06	.10	-	-
Individual	.09	.07	-	-

[1]For IWS and CPS, includes SMSA dummy and region (4 regions for CPS). For IWS and AWS includes incentive dummy and joint effects. In CPS, uses 3-digit occupation. CPS and AWS totals include 2-digit industry.
[2]Effects of interactions with year have been excluded from AWS results.
[3]The CPS sample includes all private-sector fulltime workers between the ages of 18 and 65 with reported average hourly earnings of more than $1.75.
[4]Including only observations for occupations included in the AWS sample.

Source: Tabulations from BLS *Area* Wage Survey, BLS Industry Wage Surveys (see Groshen 1988b), and May 1977 CPS.